15.458: Financial Data Science and Computing

Workshop: Data Cleaning, Signal Analysis and Risk Management

October 5, 2018

Brain Warm-up

Question: A lily-pad double in area every second. After 60 seconds it fills the pond. How long would it take to fill half the pond ?
Answer: 59

1 Data Acquisition

Clone the Github repository:

https://github.com/ucfbrd/15-458-playground.git

2 Data Pre-Processing

The first step in this endeavor is data cleaning. Identify any errors in the **Portfolio** column, flag them with a note, and suggest a corrected value or if advisable, you may choose to ignore them for purposes of your analysis.
Hint: NA, P>0,Consistency...
Required Output: Present a Systematic approach framework to identify errors and a suggested solution to these problems.

3 Data Analysis: Signal Analysis

In this part, you are free to run any kind of analysis on the four signals that you have in the csv file. The "signals" in the table are received at the end of the day (that is to say, contemporaneous with the portfolio value on the spreadsheet). We do not have a prior conviction about the forecast horizon over which signals are effective.
Hint Linear Regression, correlation, lag, plots...
Required Output Present an analysis of the predictive power of this signal, with your conclusions about its viability and shortcomings, and your brief recommendation to a Portfolio Manager. This may take the form of a brief note with any relevant graphs.

3 Data Analysis: Risk Management

In this section we will try to get a sense of what kind of portfolio we have. First compute the returns of the portfolio value. Then generate some descriptive statistics and analyze the results. Then compute the Sharpe Ratio (Consider $r_f = 0$). What do you think ?
Now go to the Fama-French Website[1] and download factor loadings on a daily basis. Use these factors to run a linear regression and analyze your results.
Hint β,α...
Required Output Present a Systematic approach to descriptive statistics when analyzing your data. Also, have a sense of the risk exposure of your portfolio.

4 Data Presentation

Prepare a 1-min summary with plots using ggplot2 and present what you have done today by designate one person.

[1] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html